

# APPLICATION UNDER UNITED STATES PATENT LAWS

Atty. Dkt. No. PW 277176/Z70389/UST  
(M#)...

Invention: POLYMORPHISMS IN THE HUMAN VCAM-1 GENE, SUITABLE FOR DIAGNOSIS AND TREATMENT OF VCAM-1 LIGAND MEDIATED DISEASES

Inventor (s): MORTEN, John Edward Norris

Pillsbury Winthrop LLP  
Intellectual Property Group  
1100 New York Avenue, NW  
Ninth Floor  
Washington, DC 20005-3918  
Attorneys  
Telephone: (202) 861-3000

## This is a:

- ☐ Provisional Application
- ☐ Regular Utility Application
- ☐ Continuing Application
  - ☒ The contents of the parent are incorporated by reference
- ☒ PCT National Phase Application
- ☐ Design Application
- ☐ Reissue Application
- ☐ Plant Application
- ☐ Substitute Specification
  - Sub. Spec Filed \_\_\_\_\_
  - in App. No. \_\_\_\_\_ / \_\_\_\_\_
- ☐ Marked up Specification re
  - Sub. Spec. filed \_\_\_\_\_
  - In App. No. \_\_\_\_\_ / \_\_\_\_\_

## SPECIFICATION

## POLYMORPHISMS IN THE HUMAN VCAM-1 GENE, SUITABLE FOR DIAGNOSIS AND TREATMENT OF VCAM-1 LIGAND MEDIATED DISEASES

This invention relates to polymorphisms in the human Vascular Cell Adhesion Molecule-1 (VCAM-1) gene. The invention also relates to methods and materials for  
5 analysing allelic variation in the VCAM-1 gene, and to the use of VCAM-1 polymorphism in the diagnosis and treatment of VCAM-1 ligand mediated diseases such as multiple sclerosis, rheumatoid arthritis, atherosclerosis and allergic asthma.

VCAM-1, also known as CD106, is a 90-110 kDa glycoprotein member of the immunoglobulin superfamily expressed mainly on the surface of activated vascular endothelial  
10 cells. Two forms of human VCAM-1 have been identified, a predominant form containing seven immunoglobulin domains and an alternatively-spliced form missing the fourth immunoglobulin domain. VCAM-1 is also found as a soluble form in serum, probably as a result of proteolytic cleavage of endothelial cell surface VCAM-1. Cell adhesion molecules have been reviewed in Mousa *et al.* (1997), DDT, 2, 187-199.

15 VCAM-1 is a ligand for the  $\alpha_4$  integrins,  $\alpha_4\beta_1$ , also known as Very Late Antigen-4 (VLA-4) or CD49d/CD29, and  $\alpha_4\beta_7$ . These integrins are members of a family of heterodimeric cell surface receptors that are composed of non-covalently associated glycoprotein subunits ( $\alpha$  and  $\beta$ ) and are involved in the adhesion of cells to other cells or to extracellular matrix. Integrin  $\alpha_4\beta_1$  is expressed on numerous haematopoietic cells, including  
20 haematopoietic precursors, peripheral and cytotoxic T lymphocytes, B lymphocytes, monocytes, thymocytes and eosinophils. Integrin  $\alpha_4\beta_7$  is expressed on lymphocytes that preferentially home to gastrointestinal mucosa and gut-associated lymphoid tissue. The  $\alpha_4$  integrins recognise a short amino acid sequence, glutamine-isoleucine-aspartic acid-serine-proline (QIDSP), exposed on the C-D loop of immunoglobulin domains 1 and 4. An accessory  
25 binding site may also be located in the adjacent immunoglobulin domain.

Expression of VCAM-1 on unactivated vascular endothelial cells is low or absent but is upregulated in human inflammatory diseases such as rheumatoid arthritis, multiple sclerosis, allergic asthma and atherosclerosis. Soluble VCAM-1 is elevated in serum and cerebrospinal fluid of multiple sclerosis patients and in serum during inflammatory bowel disease and after  
30 cardiac transplantation. *In vitro*, endothelial cells can be induced to express VCAM-1 by inflammatory cytokines such as tumour necrosis factor and interleukin-1 or by oxidative stress. VCAM-1 gene expression in vascular endothelial cells during inflammation is regulated by

transcriptional activation, thought to involve the dimeric transcription factor, nuclear factor- $\kappa$ B (NF- $\kappa$ B).

The activation and extravasation of blood leukocytes plays a major role in the development and progression of inflammatory diseases. Cell adhesion to the vascular endothelium is required before cells migrate from the blood into inflamed tissue and is mediated by specific interactions between cell adhesion molecules on the surface of vascular endothelial cells and circulating leukocytes.  $\alpha_4$  integrin/VCAM-1 binding is believed to have an important role in the recruitment of lymphocytes, monocytes and eosinophils during inflammation. Monoclonal antibodies directed against the  $\alpha_4$  integrin subunit have been shown to be effective in a number of animal models of human inflammatory diseases including multiple sclerosis, rheumatoid arthritis, allergic asthma, contact dermatitis, transplant rejection, insulin-dependent diabetes, inflammatory bowel disease, and glomerulonephritis.

$\alpha_4\beta_1$  /VCAM-1 binding has also been implicated in T-cell proliferation, B-cell localisation to germinal centres, haematopoietic progenitor cell localisation in the bone marrow, angiogenesis, placental development, muscle development and tumour cell metastasis.

Small molecule inhibitors of VCAM-1 binding to  $\alpha_4$  integrins have been designed based on the QIDSP motif in VCAM-1 and similar motifs in other  $\alpha_4$  integrin ligands fibronectin and mucosal addressin cell adhesion molecule-1 (MAdCAM-1). Small molecule and monoclonal antibody inhibitors of VCAM-1 binding to  $\alpha_4$  integrins and inhibitors of VCAM-1 expression may have utility in the treatment of autoimmune, allergic and vascular inflammatory diseases, the prevention of tumour metastasis and in mobilisation of haematopoietic progenitor cells from bone marrow prior to tumour chemotherapy.

Exon 1 of the VCAM-1 gene has been cloned and published as a EMBL Accession number: M92431 (2396 bp) and all positions herein relate to the position therein unless stated otherwise or apparent from the context.

One approach is to use knowledge of polymorphisms to help identify patients most suited to therapy with particular pharmaceutical agents (this is often termed "pharmacogenetics"). Pharmacogenetics can also be used in pharmaceutical research to assist the drug selection process. Polymorphisms are used in mapping the human genome and to elucidate the genetic component of diseases. The reader is directed to the following references for background details on pharmacogenetics and other uses of polymorphism detection:

Linder *et al.* (1997), -Clinical Chemistry, **43**, 254; Marshall (1997), Nature Biotechnology, **15**,

1249; International Patent Application WO 97/40462, Spectra Biomedical; and Schafer *et al.* (1998), Nature Biotechnology, **16**, 33.

Clinical trials have shown that patient response to treatment with pharmaceuticals is often heterogeneous. Thus there is a need for improved approaches to pharmaceutical agent  
5 design and therapy.

The present invention is based on the discovery of six single nucleotide polymorphisms (SNPs) in the VCAM-1 gene.

According to one aspect of the present invention there is provided a method for the diagnosis of a single nucleotide polymorphism in VCAM-1 in a human, which method  
10 comprises determining the sequence of the nucleic acid of the human at one or more of positions 278, 647, 707, 748, 829 and 1467 in the VCAM-1 gene as defined by the positions in EMBL ACCESSION NO. M92431, and determining the status of the human by reference to polymorphism in the VCAM-1 gene.

The term human includes both a human having or suspected of having a VCAM-1  
15 ligand mediated disease and an asymptomatic human who may be tested for predisposition or susceptibility to such disease. At each position the human may be homozygous for an allele or the human may be a heterozygote.

In one embodiment of the invention preferably the method for diagnosis described herein is one in which the single nucleotide polymorphism at position 278 is presence of T  
20 and/or C.

In another embodiment of the invention preferably the method for diagnosis described herein is one in which the single nucleotide polymorphism at position 647 is presence of A and/or G.

In another embodiment of the invention preferably the method for diagnosis described  
25 herein is one in which the single nucleotide polymorphism at position 707 is presence of T and/or C.

In another embodiment of the invention preferably the method for diagnosis described herein is one in which the single nucleotide polymorphism at position 748 is presence of T and/or C.

30 In another embodiment of the invention preferably the method for diagnosis described herein is one in which the single nucleotide polymorphism at position 829 is presence of G and/or A.

In another embodiment of the invention preferably the method for diagnosis described herein is one in which the single nucleotide polymorphism at position 1467 is presence of T and/or C.

The method for diagnosis is preferably one in which the sequence is determined by a method selected from amplification refractory mutation system and restriction fragment length polymorphism.

In another aspect of the invention we provide a method for the diagnosis of VCAM-1 ligand-mediated disease, which method comprises:

- i) obtaining sample nucleic acid from an individual,
- 10 ii) detecting the presence or absence of a variant nucleotide at one or more of positions 278, 647, 707, 748, 829 and 1467 (as defined by the position in EMBL accession number M92431), in the VCAM-1 gene and
- iii) determining the status of the individual by reference to polymorphism in the VCAM-1 gene.

Allelic variation at position 278 consists of a single base substitution from T (the published base), preferably to C. Allelic variation at position 647 consists of a single base substitution from A (the published base), preferably to G. Allelic variation at position 707 consists of a single base substitution from T (the published base), preferably to C. Allelic variation at position 748 consists of a single base substitution from T (the published base), preferably to C. Allelic variation at position 829 consists of a single base substitution from G (the published base), preferably to A. Allelic variation at position 1467 consists of a single base substitution from T (the published base), preferably to C. The status of the individual may be determined by reference to allelic variation at any one, two, three, four, five or all six positions.

The test sample of nucleic acid is conveniently a sample of blood, bronchoalveolar lavage fluid, sputum, or other body fluid or tissue obtained from an individual. It will be appreciated that the test sample may equally be a nucleic acid sequence corresponding to the sequence in the test sample, that is to say that all or a part of the region in the sample nucleic acid may firstly be amplified using any convenient technique e.g. PCR, before analysis of allelic variation.

It will be apparent to the person skilled in the art that there are a large number of analytical procedures which may be used to detect the presence or absence of variant nucleotides at one or more polymorphic positions of the invention. In general, the detection of allelic variation requires a mutation discrimination technique, optionally an amplification

reaction and optionally a signal generation system. Table 1 lists a number of mutation detection techniques, some based on the PCR. These may be used in combination with a number of signal generation systems, a selection of which is listed in Table 2. Further amplification techniques are listed in Table 3. Many current methods for the detection of allelic variation are reviewed by Nollau *et al.*, Clin. Chem. **43**, 1114-1120, 1997; and in standard textbooks, for example "Laboratory Protocols for Mutation Detection", Ed. by U. Landegren, Oxford University Press, 1996 and "PCR", 2<sup>nd</sup> Edition by Newton & Graham, BIOS Scientific Publishers Limited, 1997.

#### Abbreviations:

ALEX™	Amplification refractory mutation system linear extension
APEX	Arrayed primer extension
ARMST™	Amplification refractory mutation system
b-DNA	Branched DNA
CMC	Chemical mismatch cleavage
bp	base pair
COPS	Competitive oligonucleotide priming system
DGGE	Denaturing gradient gel electrophoresis
FRET	Fluorescence resonance energy transfer
LCR	Ligase chain reaction
MAdCAM-1	mucosal addressin cell adhesion molecule-1
MASDA	Multiple allele specific diagnostic assay
NASBA	Nucleic acid sequence based amplification
OLA	Oligonucleotide ligation assay
PCR	Polymerase chain reaction
PTT	Protein truncation test
RFLP	Restriction fragment length polymorphism
SDA	Strand displacement amplification
SNP	Single nucleotide polymorphism
SSCP	Single-strand conformation polymorphism analysis
SSR	Self sustained replication
TGGE	Temperature gradient gel electrophoresis
VCAM-1	Vascular Cell Adhesion Molecule-1
VLA-4	Very Late Antigen-4

10 Table 1 - Mutation Detection Techniques

**General:** DNA sequencing, Sequencing by hybridisation

**Scanning:** PTT\*, SSCP, DGGE, TGGE, Cleavase, Heteroduplex analysis, CMC, Enzymatic mismatch cleavage

\* Note: not useful for detection of promoter polymorphisms.

#### 15 Hybridisation Based

Solid phase hybridisation: Dot blots, MASDA, Reverse dot blots, Oligonucleotide arrays (DNA Chips)

Solution phase hybridisation: Taqman™ - US-5210015 & US-5487972 (Hoffmann-La Roche), Molecular Beacons - Tyagi *et al* (1996), Nature Biotechnology, 14, 303; WO 5 95/13399 (Public Health Inst., New York)

**Extension Based:** ARMST™, ALEX™ - European Patent No. EP 332435 B1 (Zeneca Limited), COPS - Gibbs *et al* (1989), Nucleic Acids Research, 17, 2347.

**Incorporation Based:** Mini-sequencing, APEX

**Restriction Enzyme Based:** RFLP, Restriction site generating PCR

10 **Ligation Based:** OLA

**Other:** Invader assay

Table 2 - Signal Generation or Detection Systems

**Fluorescence:** FRET, Fluorescence quenching, Fluorescence polarisation - United Kingdom Patent No. 2228998 (Zeneca Limited)

15 **Other:** Chemiluminescence, Electrochemiluminescence, Raman, Radioactivity, Colorimetric, Hybridisation protection assay, Mass spectrometry

Table 3 - Further Amplification Methods

SSR, NASBA, LCR, SDA, b-DNA

Preferred mutation detection techniques include ARMST™, ALEX™, COPS, Taqman, 20 Molecular Beacons, RFLP, and restriction site based PCR and FRET techniques.

Particularly preferred methods include ARMST™ and RFLP based methods. ARMST™ is an especially preferred method.

In a further aspect, the diagnostic methods of the invention are used to assess the efficacy of therapeutic compounds in the treatment of VCAM-1 ligand mediated diseases such 25 as autoimmune, allergic and vascular inflammatory diseases. The polymorphisms identified in the present invention occur in the promoter region of the VCAM-1 gene. The changes are not expected to alter the amino acid sequence of VCAM-1, but several of the polymorphisms affect transcription sites within the promoter region and thus may affect the transcription of the VCAM-1 gene. For example the changing of the nucleotide at position 748 (as defined by the 30 position in EMBL ACCESSION NO. M92431) from T to C results in the gain of a E1a-F rev site and the loss of a TATA box.

Assays, for example reporter-based assays, may be devised to detect whether one or more of the above polymorphisms affect transcription levels and/or message stability.

Individuals who carry particular allelic variants of the VCAM-1 gene may therefore exhibit differences in their ability to regulate protein biosynthesis under different physiological  
5 conditions and will display altered abilities to react to different diseases. In addition, differences in protein regulation arising as a result of allelic variation may have a direct effect on the response of an individual to drug therapy. The diagnostic methods of the invention may be useful both to predict the clinical response to such agents and to determine therapeutic dose.

10 In a further aspect, the diagnostic methods of the invention, are used to assess the predisposition and/or susceptibility of an individual to diseases mediated by VCAM-1 ligands. This may be particularly relevant in the development of autoimmune, allergic and vascular inflammatory diseases and other diseases which are modulated by VCAM-1 interactions. The present invention may be used to recognise individuals who are particularly at risk from  
15 developing these conditions.

In a further aspect, the diagnostic methods of the invention are used in the development of new drug therapies which selectively target one or more allelic variants of the VCAM-1 gene. Identification of a link between a particular allelic variant and predisposition to disease development or response to drug therapy may have a significant impact on the design of new  
20 drugs. Drugs may be designed to regulate the biological activity of variants implicated in the disease process whilst minimising effects on other variants.

In a further diagnostic aspect of the invention the presence or absence of variant nucleotides is detected by reference to the loss or gain of, optionally engineered, sites recognised by restriction enzymes. In the accompanying Example 1 we provide details of  
25 convenient engineered restriction enzyme sites that are lost or gained as a result of a polymorphism of the invention.

According to another aspect of the present invention there is provided a nucleic acid comprising any one of the following polymorphisms:  
the nucleic acid of EMBL ACCESSION No. M92431 with C at position 278 in the promoter  
30 sequence as defined by the position in EMBL ACCESSION No. M92431;  
the nucleic acid of EMBL ACCESSION No. M92431 with G at position 647 in the promoter sequence as defined by the position in EMBL ACCESSION No. M92431;



- the nucleic acid of EMBL ACCESSION No. M92431 with C at position 707 in the promoter sequence as defined by the position in EMBL ACCESSION No. M92431;
- the nucleic acid of EMBL ACCESSION No. M92431 with C at position 748 in the promoter sequence as defined by the position in EMBL ACCESSION No. M92431;
- 5 the nucleic acid of EMBL ACCESSION No. M92431 with A at position 829 in the promoter sequence as defined by the position in EMBL ACCESSION No. M92431;
- the nucleic acid of EMBL ACCESSION No. M92431 with C at position 1467 in the promoter sequence as defined by the position in EMBL ACCESSION No. M92431;
- or a complementary strand thereof or a fragment thereof of at least 20 bases comprising at
- 10 least one polymorphism.

Fragments are at least 17 bases, more preferably at least 20 bases, more preferably at least 30 bases. The nucleic acid of the invention does not encompass naturally occurring nucleic acid as it occurs in nature, for example, the nucleic acid is at least partially purified from at least one component with which it occurs naturally. Preferably the nucleic acid is at

15 least 30% pure, more preferably at least 60% pure, more preferably at least 90% pure, more preferably at least 95% pure, and more preferably at least 99% pure.

According to another aspect of the invention there is provided use of a nucleic acid sequence comprising at least one of the polymorphisms in the promoter disclosed herein to identify compounds that modify expression of the VCAM-1 gene. Modification of expression

20 includes inhibition or enhancement of expression. This is conveniently done by measuring expression levels of a reporter gene (for example beta-galactosidase) under the control of the promoter in transfected host cells in the presence or absence of test compounds. Suitable test compounds include polynucleotides capable of binding to the promoter through triplex strand formation. Accordingly suitable compounds can be identified for therapeutic use which alter

25 native gene expression either up or down as appropriate for the relevant disease to be treated. The reader is directed to the following references on nucleic acid triplex formation and uses: Progress in developments of Triplex-Based strategies: Giovannangeli C; Helene C: Antisense and Nucleic Acid Drug Development / 7/4 (413-421) /1997; Recent developments in triple-helix regulation of gene expression: Neidle S: Anti-Cancer Drug Design / 12/5 (433-442)

30 /1997; Triplex DNA: Fundamentals, advances, and potential applications for gene therapy: Chan PP; Glazer PM: Journal of Molecular Medicine / 75/4 (267-282) /1997; Oligonucleotide directed triple helix formation: Sun J-S; Garestier T; Helene C: Current Opinion in Structural

Biology / 6/3 (327-333) /1996; C Mayfield, M Squibb, D Miller (1994) Inhibition of nuclear protein binding to the human Ki-ras promoter by triplex-forming oligonucleotides Biochemistry 33,3358-3363; WM Olivas, LJ Maher (1996) Binding of DNA oligonucleotides to sequences in the promoter of the human bcl-2 gene Nucleic Acids Research 24, 1758-1764; 5 C Mayfield, S Ebinghaus, J Gees, D Jones, B Rodu, M Squibb, D Miller (1994) Triplex formation by the human HA-ras promoter inhibits Sp1 binding and in vitro transcription J Biol Chem 269,18232-18238; and JE Gee, GR Revankar, TS Rao, ME Hogan (1995) Triplex formation at the rat neu gene utilizing imidazole and 2'-deoxy-6-thioguanosine base substitutions Biochemistry 34,2042-2048.

- 10 According to another aspect of the present invention there is provided a computer readable medium comprising at least one novel polynucleotide sequence of the invention stored on the medium. The computer readable medium may be used, for example, in homology searching, mapping, haplotyping, genotyping or pharmacogenetic analysis or any other bioinformatic analysis. The reader is referred to Bioinformatics, A practical guide to the 15 analysis of genes and proteins, Edited by A D Baxevanis & B F F Ouellette, John Wiley & Sons, 1988. Any computer readable medium may be used, for example, compact disk, tape, floppy disk, hard drive or computer chips.

The polynucleotide sequences of the invention, or parts thereof, particularly those relating to and identifying the single nucleotide polymorphisms identified herein represent a 20 valuable information source, for example, to characterise individuals in terms of haplotype and other sub-groupings, such as investigation of susceptibility to treatment with particular drugs. These approaches are most easily facilitated by storing the sequence information in a computer readable medium and then using the information in standard bioinformatics programs or to search sequence databases using state of the art searching tools such as "GCC". Thus, the 25 polynucleotide sequences of the invention are particularly useful as components in databases useful for sequence identity and other search analyses. As used herein, storage of the sequence information in a computer readable medium and use in sequence databases in relation to 'polynucleotide or polynucleotide sequence of the invention' covers any detectable chemical or physical characteristic of a polynucleotide of the invention that may be reduced to, 30 converted into or stored in a tangible medium, such as a computer disk, preferably in a computer readable form. For example, chromatographic scan data or peak data, photographic scan or peak data, mass spectrographic data, sequence gel (or other) data.

The invention provides a computer readable medium having stored thereon one or a more polynucleotide sequences of the invention. For example, a computer readable medium is provided comprising and having stored thereon a member selected from the group consisting of: a polynucleotide comprising the sequence of a polynucleotide of the invention, a  
5 polynucleotide consisting of a polynucleotide of the invention, a polynucleotide which comprises part of a polynucleotide of the invention, which part includes at least one of the polymorphisms of the invention, a set of polynucleotide sequences wherein the set includes at least one polynucleotide sequence of the invention, a data set comprising or consisting of a polynucleotide sequence of the invention or a part thereof comprising at least one of the  
10 polymorphisms identified herein. A computer based method is also provided for performing sequence identification, said method comprising the steps of providing a polynucleotide sequence comprising a polymorphism of the invention in a computer readable medium; and comparing said polymorphism containing polynucleotide sequence to at least one other polynucleotide or polypeptide sequence to identify identity (homology), i.e. screen for the  
15 presence of a polymorphism.

The invention further provides nucleotide primers which can detect the polymorphisms of the invention.

According to another aspect of the present invention there is provided an allele specific primer capable of detecting a VCAM-1 gene polymorphism at one or more of positions 278,  
20 647, 707, 748, 829 and 1467 in the VCAM-1 gene as defined by the positions in EMBL  
ACCESSION NO. M92431.

An allele specific primer is used, generally together with a constant primer, in an amplification reaction such as a PCR reaction, which provides the discrimination between alleles through selective amplification of one allele at a particular sequence position e.g. as  
25 used for ARMS<sup>TM</sup> assays. The allele specific primer is preferably 17- 50 nucleotides, more preferably about 17-35 nucleotides, more preferably about 17-30 nucleotides.

An allele specific primer preferably corresponds exactly with the allele to be detected but derivatives thereof are also contemplated wherein about 6-8 of the nucleotides at the 3' terminus correspond with the allele to be detected and wherein up to 10, such as up to 8, 6, 4,  
30 2, or 1 of the remaining nucleotides may be varied without significantly affecting the properties of the primer.

Primers may be manufactured using any convenient method of synthesis. Examples of such methods may be found in standard textbooks, for example "Protocols for Oligonucleotides and Analogues; Synthesis and Properties," Methods in Molecular Biology Series; Volume 20; Ed. Sudhir Agrawal, Humana ISBN: 0-89603-247-7; 1993; 1<sup>st</sup> Edition. If  
5 required the primer(s) may be labelled to facilitate detection.

According to another aspect of the present invention there is provided an allele-specific oligonucleotide probe capable of detecting a VCAM-1 gene polymorphism at one or more of positions 278, 647, 707, 748, 829 and 1467 in the VCAM-1 gene as defined by the positions in EMBL ACCESSION NO. M92431.

10 The allele-specific oligonucleotide probe is preferably 17- 50 nucleotides, more preferably about 17-35 nucleotides, more preferably about 17-30 nucleotides.

The design of such probes will be apparent to the molecular biologist of ordinary skill. Such probes are of any convenient length such as up to 50 bases, up to 40 bases, more conveniently up to 30 bases in length, such as for example 8-25 or 8-15 bases in length. In  
15 general such probes will comprise base sequences entirely complementary to the corresponding wild type or variant locus in the gene. However, if required one or more mismatches may be introduced, provided that the discriminatory power of the oligonucleotide probe is not unduly affected. The probes of the invention may carry one or more labels to facilitate detection.

According to another aspect of the present invention there is provided a diagnostic kit  
20 comprising an allele specific oligonucleotide probe of the invention and/or an allele-specific primer of the invention.

The diagnostic kits may comprise appropriate packaging and instructions for use in the methods of the invention. Such kits may further comprise appropriate buffer(s) and polymerase(s) such as thermostable polymerases, for example taq polymerase.

25 In another aspect of the invention, the single nucleotide polymorphisms of this invention may be used as genetic markers in linkage studies. This particularly applies to the polymorphism at 278 (as defined by the position in EMBL ACCESSION NO. M92431) because of its relatively high frequency (see below). The VCAM-1 gene has been mapped to chromosome 1p31-32 (Cybulsky et al Proc. Natl. Acad. Sci. USA 88, 7859-7863, 1991).

30 Low frequency polymorphisms may be particularly useful for haplotyping as described below. A haplotype is a set of alleles found at linked polymorphic sites (such as within a gene) on a single (paternal or maternal) chromosome. If recombination within the gene is random,

there may be as many as  $2^n$  haplotypes, where 2 is the number of alleles at each SNP and n is the number of SNPs. One approach to identifying mutations or polymorphisms which are correlated with clinical response is to carry out an association study using all the haplotypes that can be identified in the population of interest. The frequency of each haplotype is limited by the frequency of its rarest allele, so that SNPs with low frequency alleles are particularly useful as markers of low frequency haplotypes. As particular mutations or polymorphisms associated with certain clinical features, such as adverse or abnormal events, are likely to be of low frequency within the population, low frequency SNPs may be particularly useful in identifying these mutations (for examples see: Linkage disequilibrium at the cystathionine beta synthase (CBS) locus and the association between genetic variation at the CBS locus and plasma levels of homocysteine. *Ann Hum Genet* (1998) 62:481-90, De Stefano V, Dekou V, Nicaud V, Chasse JF, London J, Stansbie D, Humphries SE, and Gudnason V; and Variation at the von willebrand factor (vWF) gene locus is associated with plasma vWF:Ag levels: identification of three novel single nucleotide polymorphisms in the vWF gene promoter. *Blood* (1999) 93:4277-83, Keightley AM, Lam YM, Brady JN, Cameron CL, Lillicrap D).

According to another aspect of the present invention there is provided a method of treating a human in need of treatment with a VCAM-1 ligand antagonist drug in which the method comprises:

- i) diagnosis of a single nucleotide polymorphism in VCAM-1 gene in the human, which diagnosis comprises determining the sequence of the nucleic acid at one or more of positions 278, 647, 707, 748, 829 and 1467 (as defined by the position in EMBL accession number M92431), and determining the status of the human by reference to polymorphism in the VCAM-1 gene; and
  - ii) administering an effective amount of a VCAM-1 ligand antagonist.
- Preferably determination of the status of the human is clinically useful. Examples of clinical usefulness include deciding which antagonist drug or drugs to administer and/or in deciding on the effective amount of the drug or drugs.

VCAM-1 ligand antagonist drugs have been disclosed in the following publications: international patent application WO 97/49731, Zeneca Limited; international patent application WO 97/02289, Zeneca Limited; international patent application WO 96/20216, Zeneca Limited; US patent 5510332, Texas Biotechnology; international patent application WO 96/01644, Athena Neurosciences; international patent application WO 96/01644, Athena

Neurosciences and; international patent application WO 96/00581, Zeneca Limited. A VCAM-1 ligand antagonist drug may act directly at VCAM-1 and/or at a ligand, such as VLA-4, which binds to VCAM-1.

According to another aspect of the present invention there is provided use of a VCAM-1 ligand antagonist drug in preparation of a medicament for treating a VCAM-1 mediated disease in a human diagnosed as having a single nucleotide polymorphism at one or more of positions 278, 647, 707, 748, 829 and 1467 (as defined by the position in EMBL accession number M92431).

According to another aspect of the present invention there is provided a pharmaceutical pack comprising VCAM-1 ligand antagonist drug and instructions for administration of the drug to humans diagnostically tested for a single nucleotide polymorphism at one or more of positions 278, 647, 707, 748, 829 and 1467 (as defined by the position in EMBL accession number M92431).

The invention will now be illustrated but not limited by reference to the following Examples. All temperatures are in degrees Celsius.

In the Examples below, unless otherwise stated, the following methodology and materials have been applied.

AMPLITAQ™, available from Perkin-Elmer Cetus, is used as the source of thermostable DNA polymerase.

General molecular biology procedures can be followed from any of the methods described in "Molecular Cloning - A Laboratory Manual" Second Edition, Sambrook, Fritsch and Maniatis (Cold Spring Harbor Laboratory, 1989).

Electropherograms were obtained in a standard manner: data was collected by ABI377 data collection software and the wave form generated by ABI Prism sequencing analysis (2.1.2).

#### Example 1

#### **Identification of Polymorphisms**

##### **1. Methods**

##### DNA Preparation

DNA was prepared from frozen blood samples collected in EDTA following protocol I (Molecular Cloning: A Laboratory Manual, p392, Sambrook, Fritsch and Maniatis, 2<sup>nd</sup> Edition, Cold Spring Harbor Press, 1989) with the following modifications. The thawed blood was

diluted in an equal volume of standard saline citrate instead of phosphate buffered saline to remove lysed red blood cells. Samples were extracted with phenol, then phenol/chloroform and then chloroform rather than with three phenol extractions. The DNA was dissolved in deionised water.

## 5 Template Preparation

Templates were prepared by PCR using the oligonucleotide primers and annealing temperatures set out below. The extension temperature was 72° and denaturation temperature 94°, each step was 1 minute. All reactions contained 1 mM MgCl<sub>2</sub>. Generally 50 ng of genomic DNA was used in each reaction and subjected to 35 cycles of PCR.

Fragment	Forward Oligo	Reverse Oligo	Annealing Temp
22-1151	22-41	1131-1151	62°
595-1151	595-618	1131-1151	58°
1264-1804	1264-1286	1782-1804	58°

10

For dye-primer sequencing the forward primers were modified to include M13 forward sequence (ABI protocol P/N 402114, Applied Biosystems) at the 5' end of the oligonucleotide. DNA polymerase (Amplitaq Gold™, Perkin Elmer Cetus) was used to generate products 595-1151 and 1264-1804.

## 15 Dye Primer Sequencing

Dye-primer sequencing using M13 forward and reverse primers was as described in the ABI protocol P/N 402114 for the ABI Prism™ dye primer cycle sequencing core kit with "AmpliTaq FS"™ DNA polymerase, modified in that the annealing temperature was 45° and DMSO was added to the cycle sequencing mix to a final concentration of 5 %.

20 The extension reactions for each base were pooled, ethanol/sodium acetate precipitated, washed and resuspended in formamide loading buffer.

4.25 % Acrylamide gels were run on an automated sequencer (ABI 377, Applied Biosystems).

## 2. Results

## 25 Novel Polymorphisms

Position <sup>1</sup>	Published <sup>2</sup>	Variant	RFLP	Variant Allele Frequency	TF <sup>3</sup> Site Gain	TF <sup>3</sup> site Loss

278	T	C	loss of Vsp I	41/94	LF-A2 rev, HNF1 rev, SBF-1 rev, phyA3 rev	AP-1 rev
647	A	G	engineered Pvu I (see Example 2)	1/82	none	HNF-5, ZRE 3, 4,&6, GR intron site 4
707	T	C	none	9/82	none	none
748	T	C	gain of Bst F5 I	2/82	E1a-F rev	TATA box
829	G	A	gain of Ksp 632 I	1/108	none	MalT
1467	T	C	gain of Rsa I	1/82	none	TEF- 1,EF1L,S ph box, D10

<sup>1</sup>As defined by the position in EMBL ACCESSION NO. M92431

<sup>2</sup>Iadernarco *et al.* J. Biol. Chem, 267, 16323-16329, 1992.

<sup>3</sup>TF = transcription factor

Allele frequency was determined in a European control population.

5

### Example 2

#### **Engineered restriction sites for detection of polymorphisms**

Standard methodology can be used to detect the polymorphism at positions 647 (as  
10 defined by the position in EMBL ACCESSION NO. M92431) based on the materials set out  
below.

Diagnostic Fragment	Forward primer	Reverse primer
518-669	518-540	648-669 Pvu I

Reverse primer sequence CCCAGAGGTCCTTTACAGCGAT (SEQ ID NO:1).

The product generated by these primers will include a Pvu I site, only if the diagnostic  
15 fragment contained a G allele at position 647.

#### **Sequence Listing Free Text**



<223> Description of Artificial Sequence:PCR primer